

IN THE UNITED STATES  
PATENT AND TRADEMARK OFFICE



Attorney Docket No.: GOOGLE-5

Applicants: William PUGH, Monika H. HENZINGER

Serial No.: Not yet assigned

Filing Date: Filed herewith

Title: DETECTING DUPLICATE AND NEAR-DUPLICATE FILES

Examiner: Not yet assigned

Group Art Unit: Not yet assigned

Assistant Commissioner for Patents  
**Box Patent Application**  
Washington, D. C. 20231

S I R:

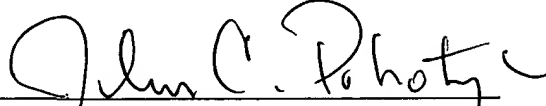
Information Disclosure Statement

The applicants respectfully request that the following references listed on the attached PTO-FORM-1449 be considered in the examination of the above-identified application. A copy of each of these references is enclosed.

Applicants preserve the right to establish that any of the references listed on the attached PTO-FORM-1449 are not prior art to the above-captioned application.

Respectfully submitted,

Dated: January 24, 2001


  
John C. Pokotylo, Attorney  
Reg. No. 36,242  
Customer No. 26479  
(732) 335-1222

STRAUB & POKOTYLO  
1 Bethany Road  
Suite 83, Bldg. 6  
Hazlet, New Jersey 07730

\*\*\*EXPRESS MAIL CERTIFICATION\*\*\*

"Express Mail" mailing label number: EL667251055US  
Date of deposit: January 24, 2001

I hereby certify that this paper or fee is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above and is addressed to the Assistant Commissioner of Patents, **Box Patent Application**, Washington, D.C. 20231.

  
\_\_\_\_\_  
John C. Pokotylo

\_\_\_\_\_  
36,242  
Reg. No.

U.S. Department of Commerce Patent and Trademark Office			<i>Complete if Known</i>		
LIST OF REFERENCES CITED BY APPLICANT (use as many sheets as necessary)			Application Number: <b>Not yet assigned</b> Filing Date: <b>Herewith</b> First Named Inventor: <b>William PUGH</b> Group Art Unit: <b>Not yet assigned</b> Examiner Name: <b>Not yet assigned</b>		
Sheet	1	of	1	Attorney Docket No.: <b>GOOGLE-5</b>	

OTHER REFERENCES - NON-PATENT LITERATURE DOCUMENTS				
Examiner Initials*	Cite No. <sup>1</sup>	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume, issue number(s), publisher, country, where published, source		T <sup>2</sup>
	AA.	Andrei Z. Broder, Steven C. Glassman, Mark S. Manasse, Geoffrey Zweig, "Syntactic Clustering of the Web", <u>Proc. 6<sup>th</sup> Int'l. World Wide Web Conf.</u> , pp. 391-404 (April 1997) and <u>SRC Technical Note 1997-015</u> (July 25, 1997) (Downloaded from <a href="http://gatekeeper.dec.com/pub/DEC/SRC/technical-notes/SRC-1997-015-html/">http://gatekeeper.dec.com/pub/DEC/SRC/technical-notes/SRC-1997-015-html/</a> on November 27, 2000, pp. 1-14.)		
	AB.	Andrei Z. Broder, "On the resemblance and containment of documents", <u>Proc. of Compression and Complexity of Sequences 1997</u> , pp. 21-29 (IEEE Computer Society) pp. 1-9.		
	AC.	Sergey Brin, James Davis, Hector Garcia-Molina, "Copy Detection Mechanisms for Digital Documents", <u>Proc. of ACM SIGMOD Annual Conf.</u> (San Jose, 1995) (Downloaded from <a href="http://www-db.stanford.edu/~sergey/copy.html">http://www-db.stanford.edu/~sergey/copy.html</a> on November 27, 2000, pp. 1-21.)		
	AD.	Andrei Z. Broder, "Some applications of Rabin's fingerprinting method", R. Capocelli, A. DeSantis, U. Vaccaro, Eds; <u>Sequences II: Methods in Communications, Security, and Computer Science</u> , pp. 143-152 (Springer-Verlag, 1993).		
	AE.	Min Fang, Narayanan Shivakumar, Hector Garcia-Molina, Rajeev Motwani, Jeffrey D. Ullman, "Computing Iceberg Queries Efficiently", <u>Proc. 24<sup>th</sup> Int'l. Conf. on Very Large Databases</u> , pp. 299-310 (1998).		

Examiner Signature		Date Considered	
--------------------	--	-----------------	--

\*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

1 Unique citation designation number. 2 Applicant is to place a check mark here if English language translation is attached.